

METHOD AND APPARATUS FOR CHECKING CONTINUITY OF LEAF-TO-ROOT VLAN CONNECTIONS

Technical Field

- 5 [0001] The invention relates to data communication networks. The invention relates specifically to checking continuity of connections made across a connection-based network. The connection-based network could be a cell relay network such as an ATM network or a multi protocol label switching (MPLS) network. The invention has application in providing fault tolerance in
- 10 virtual LANs (VLANs) in which data is carried over a connection-based network.

Background

- 15 [0002] In this disclosure the term virtual LAN ("VLAN") means a data communication network which comprises a number of segments which are connected together by links which operate according to a networking protocol different from that of the network segments. For example, two segments of an ethernet network may be connected by way of a connection in a connection-based network. The connection may be, for example, a virtual circuit in an
- 20 asynchronous transfer mode (ATM) network or a path in an MPLS network. Each of the network segments may be interfaced to the connection-based network by a bridge.
- 25 [0003] Data may be delivered between the segments in any of a wide number of ways. For example, where the segments comprise ethernet segments and the network which connects the segments comprises an ATM network, then ethernet frames may be carried between the segments according to standards such as local area network emulation ("LANE"), multi-protocol over ATM ("MPOA") or IP over ATM. The particular format by which data is conveyed
- 30 between network segments is not important to this invention.

[0004] The topology of a VLAN is typically initially established using a protocol such as the spanning tree protocol (STP). STP generates a network topology which is defined by a spanning tree. The spanning tree defines a topology which does not include any loops. After the topology has been established, a bridge at a root of the spanning tree generates and sends toward the leaves of the spanning tree bridge protocol data units BPDUs. If a bridge at a leaf in the spanning tree fails to receive a BPDU within a defined time period, the bridge can trigger a request for a topology change. If a new path to the bridge exists then the topology change will incorporate the new path into the network.

[0005] IEEE standard 802.1Q provides a set of capabilities which permit media access control (MAC) bridges to define and manage networks in which multiple broadcast domains can co-exist on a single physical medium. This IEEE standard uses the term "VLAN" to describe such networks. In this disclosure the term VLAN is not used in the same sense as it is in IEEE 802.1Q (although this invention may be applied to networks which operate, or which include portions which operate, according to IEEE 802.1Q). IEEE standard 802.1D describes the operation of MAC bridges.

[0006] There are various standards applicable to ATM networks in which operations, administration, and management ("OAM") cells are used for monitoring connections within the ATM network. These include Bellcore UNI specification 3.1 and ITU-I.610.

[0007] U.S. patent No. 6,181,680 describes a method for testing ATM networks using OAM cells.

[0008] U.S. patent No. 5,901,141 discloses a method for identifying those nodes passed through by a virtual connection which joins two end nodes. The method sends trace cells downstream from a monitoring node toward an endpoint. Each downstream node either loops back a trace cell identifying itself to the monitoring node or forwards the trace cell further.

[0009] There exists a need for cost-effective methods and apparatus for providing fault tolerance mechanisms in VLANs.

10 Summary of the Invention

[0010] The invention provides methods and apparatus for monitoring the continuity of the connections in a connection-based network which connect plural segments of a virtual LAN. One aspect of the invention provides a method for providing fault tolerance in a VLAN having a topology defined by a spanning tree having a root node and at least one leaf node. The root and leaf nodes are interconnected by connections in a connection-based network. The method comprises: sending from a first node in a connection used by the VLAN, in a leaf-to-root direction a series of continuity checking cells; detecting the continuity checking cells at a second node in the connection wherein the second node is located between the first node and the root node; and, generating a request for a change in the topology of the VLAN in response to not receiving one or more continuity checking cells at the second node.

[0011] The method may include generating a spanning tree protocol topology change notification. In some embodiments the request for a change in the topology of the VLAN may result in the generation of a connection rerouting request.

[0012] The segments may comprise ethernet segments. The connection-based network may comprise an ATM network. The continuity checking cells may comprise OAM cells.

- 5 [0013] Another aspect of the invention provides a method for rerouting a connection in a connection-based network where the connection carries data traffic between ethernet segments of a VLAN. The method comprises: configuring nodes at first and second ends of the connection respectively to source and sink continuity checking cells; sending continuity checking cells at a specified rate from the node at the first end of the connection; receiving the continuity checking cells at a cell sink at the node at the second end of the connection; generating a request for a change in topology of the VLAN in response to the cell sink not receiving a predetermined number of the continuity checking cells; generating a reroute signal for the connection in response to the request for a change in topology of the VLAN and, rerouting the connection through the connection-based network in response to the reroute signal.

- [0014] In some embodiments, determining that the cell sink has not received a predetermined number of the continuity checking cells comprises determining that a time elapsed since receipt of a most recently received one of the continuity checking cells exceeds a threshold time.

- [0015] A further aspect of the invention provides a method for providing fault tolerance in an ethernet VLAN comprising a plurality of ethernet segments connected to an ATM network by bridges and an ATM virtual circuit extending between a first one of the bridges and a second one of the bridges. The method comprises: configuring nodes at first and second ends of the virtual circuit respectively to source and sink OAM continuity checking cells; sending from a source port at the first end of the virtual circuit OAM continuity

checking cells at a specified rate; receiving the OAM continuity checking cells at a sink port at the second end of the virtual circuit; and, generating an action request in response to the sink port determining that it has not received a number of the OAM continuity checking cells.

5

[0016] Another aspect of the invention provides a virtual LAN having a topology defined by a spanning tree. The virtual LAN comprises: a plurality of network segments each bridged to a connection-oriented network; a plurality of connections in the connection-based network, the connections interconnecting the plurality of network segments; a cell source located on a first one of the connections, the cell source configured to generate and send on the connection temporally spaced apart continuity checking cells in a direction toward a root of the spanning tree; and, a cell sink located on the first one of the connections at a location between the cell source and the root of the spanning tree. The cell sink is configured to receive the continuity checking cells and to trigger an alarm condition in response to not receiving one or more of the continuity checking cells sent by the cell source. Optionally continuity checking cells may also be sent on the first one of the connections in a direction toward a leaf of the spanning tree. The continuity checking cells may be sent in the root-to-leaf direction at a rate higher than a rate at which BPDUs are sent in the root-to-leaf direction.

[0017] Further aspects of the invention and features of specific embodiments of the invention are described below.

25

Brief Description of the Drawings

[0018] In drawings which illustrate non-limiting embodiments of the invention:

Figure 1 is a schematic diagram of a network having a number of ethernet segments connected by cell relay connections;

Figure 2 is a block diagram illustrating apparatus according to an embodiment of the invention;

5 Figure 3 is a diagram illustrating a spanning tree for a network having a very simple topology; and,

Figure 4 is a flow chart illustrating a method according to an embodiment of the invention.

10 Description

[0019] Throughout the following description, specific details are set forth in order to provide a more thorough understanding of the invention. However, the invention may be practiced without these particulars. In other instances, well known elements have not been shown or described in detail to avoid unnecessarily obscuring the invention. Accordingly, the specification and drawings are to be regarded in an illustrative, rather than a restrictive, sense.

[0020] Figure 1 illustrates a simple network 10. Network 10 has three
20 segments 12A, 12B and 12C. The segments may comprise broadcast-based networks, such as ethernet segments. Segments 12A, 12B, and 12C are connected by a connection-based network 14 to provide a VLAN. The connection-based network may comprise, for example, a cell relay network which could be an asynchronous transfer mode ("ATM") network, a multi-
25 protocol label switching ("MPLS") network, or the like. Segments 12 are each bridged to the connection-based network. Bridges 16A, 16B and 16C (collectively bridges 16) respectively provide interfaces between ethernet segments 12A, 12B and 12C and connection-based network 14.

[0021] The VLAN may include a mechanism for providing fault-tolerance at the VLAN level. For example, a bridge **16** at a root of a spanning tree which defines a topology for network **10** may send BPDUs to bridges **16** at leaves of the spanning tree. Upon failing to receive an expected BPDUs the
5 bridge **16** at a leaf of the spanning tree may trigger a network reconfiguration. The network reconfiguration may be triggered by the bridge **16** sending a topology change notification to another bridge **16** at the root of the spanning tree.

[0022] The connection-based network may comprise a cell relay network, as illustrated in Figure 1. Cell relay network **14** may comprise, for example, an ATM network.

[0023] As shown in Figure 2, in one embodiment of the invention, each
15 bridge **16** comprises a packet source **20** and a packet sink **22**. The term "packet source" is used herein to mean a source of packets, which may comprise cells or fixed or variable length data frames of other types. The term "packet sink" is used herein to mean a sink for packets sourced at a packet source. Packet source **20** can generate a series of continuity checking (CC) packets at a
20 specified rate. In the illustrated embodiment of the invention, packet source **20** comprises a cell source, packet sink **22** comprises a cell sink and the CC packets comprise CC cells.

[0024] The CC cells may be generated at a suitable rate which may be,
25 for example, in the range of one cell per 5 seconds to one cell per 0.1 second. In certain implementations the rate may be one cell per second. A timer **21** associated with cell source **20** provides a signal used by cell source **20** to generate and send CC cells at spaced apart times. Timer **21** may be integrated with cell source **20** or may comprise a separate timer accessible to cell source

20. Cell sinks **22** and cell sources **20** may be located at nodes in cell relay network **14** other than bridges **16**.

[0025] Where the connection-based network comprises an ATM

- 5 network, the cells generated by cell source **20** may be operations, administration, and management ("OAM") cells. Cell sink **22** receives CC cells inserted by a cell source **20**.

[0026] A cell sink **22** may be configured to monitor for the CC cells

- 10 from a cell source **20**. Cell sink **22** includes or is associated with a mechanism which is configured to trigger an alarm condition if a specified number of expected CC cells are missed or, equivalently, if a specified time elapses after the detection of an CC cell without another CC cell being detected. Cell sink **22** is also associated with a timer. Where a single device hosts both a cell source
15 **20** and a cell sink **22** the same timer **21** may be used by both cell source **20** and cell sink **22**. In the alternative, separate timers may be provided.

[0027] Cell sink **22**, or another component which detects the alarm

- 20 condition, may trigger a request which results in the affected connection being bypassed. The request may comprise a request to a VLAN-level fault tolerance mechanism. The request may result in the rerouting of the affected connection (i.e. the request may result in a VC reroute signal being generated, for example, the request may cause the VLAN to generate a soft permanent virtual connection ("SPVC") reroute request). The use of SPVC reroute requests to
25 repair failed virtual circuits is discussed in commonly-owned U.S. patent application No. 09/821,708 entitled METHOD AND APPARATUS FOR REROUTING A CONNECTION IN A DATA COMMUNICATION NETWORK BASED ON A USER CONNECTION MONITORING FUNCTION which is hereby incorporated by reference. The request may also

comprise a request which results in the network topology being reconfigured (this may, for example, take the form of a STP topology change notification or equivalent, or a request that VLAN data currently being carried by way of a particular bridge port be carried instead by an auxiliary bridge port).

5

[0028] Cell sources **20** and cell sinks **22** may be implemented as software processes executed by a suitable data processor, as hardware devices, or as a combination of hardware and software. Cell sources **20** and cell sinks **22** may be implemented, for example, by circuits provided in an application specific integrated circuit ("ASIC") or field programmable gate array ("FPGA").

10

[0029] At least some of the cell sinks **22** are located at nodes in cell relay network **14** which cells reach from corresponding cell sources **20** by travelling toward the root of a spanning tree which defines the topology of a VLAN (i.e. by travelling in a leaf-to-root direction). The cell sources may be located at leaf nodes of the spanning tree and/or at other nodes along the connections through which data flows on its way from the leaf nodes to the root node of the spanning tree. The cell sinks are located on the connections between the corresponding cell source and the root of the spanning tree. In this disclosure "between the corresponding cell source and the root of the spanning tree" includes locations at the root of the spanning tree. In some embodiments of the invention the CC cells are generated at a leaf node of a spanning tree which defines a VLAN and are detected at a root node of the spanning tree.

15

20

25

[0030] It may be desirable in many cases, but is not necessary in all implementations of the invention, that the cell sources **20** and cell sinks **22** be arranged so that continuity checking packets travel over the full length of all connections included in the topology of the VLAN. The invention could be used to detect faults in a subset of the connections used by a VLAN or to detect

faults which may occur in one or more selected segments of one or more connections used by the VLAN.

[0031] Cell sources **20** and corresponding cell sinks **22** may also be
5 located at nodes in cell relay network **14** such that continuity checking cells
from the cell sources **20** travel along the connections used by the VLAN toward
leaves of the spanning tree (i.e. in a root-to-leaf direction). Such continuity
checking cells may be generated at a rate higher than any BPDUs which may be
sent from the root of the spanning tree to the leaves of the spanning tree as part
10 of a VLAN-level fault tolerance mechanism.

[0032] In the example embodiment of Figure 3, the topology of a VLAN
is defined by a spanning tree **30** having a root at bridge **16A** and leaves at
bridges **16B** and **16C**. Cell sources at bridges **16B** and **16C** source CC cells
15 which travel in leaf-to-root directions as indicated by arrows **32A** and **32B** to a
cell sink **22** (or multiple cell sinks **22**) at bridge **16A**.

[0033] The invention provides a method for monitoring continuity of
connections in a VLAN. As shown in Figure 4, the method **100** comprises, at
20 a bridge connecting a VLAN segment to a cell relay network, generating and
sending over a connection used by the VLAN toward a root node of a spanning
tree for the VLAN a series of continuity checking (CC) cells (block **104**). The
CC cells may be sent continuously at a rate in the range of, for example, once
per 1/10 second to once per 5 seconds. The rate at which the CC cells is sent is
25 preferably in the range of once per ½ second to once per 2 seconds. In a
currently preferred embodiment of the invention the CC cells are sent at a rate
of approximately once per second.

[0034] CC cells may be sent at a higher rate to permit faster fault recovery. Sending CC cells at a higher rate consumes bandwidth. For each application a tradeoff between bandwidth consumed and the maximum time for detecting a loss of connectivity in the leaf-to-root direction may be achieved by selecting a suitable rate for sending the CC cells.

[0035] In block 106 the CC cells are detected at an end of the connection. As long as the CC cells continue to be detected at the end of the connection it is known that the connection is transmitting cells. The time since receipt of the last CC cell is monitored in block 107. If the time exceeds a threshold as determined in block 108 then an alarm condition is triggered in block 110. Triggering the alarm condition may comprise generating an action request (block 112). The action request may comprise a request to a VLAN-level fault tolerance mechanism. The action request may comprise, for example, a topology change request or a request that the broken connection be rerouted.

[0036] Where the action request comprises a topology change request, the topology change request may comprise a BPDU. The BPDU may be issued to a bridge at the root node and may be of the same nature as the BPDU which would be issued by a leaf-node upon loss of connectivity to a bridge at the root node.

[0037] The actions of all of the blocks indicated by reference numeral 105 (which include block 106 and blocks 107 to 112) may be performed in a device at the location where the CC cells are detected in block 106.

[0038] Method 100 may also monitor continuity of the connection in a direction away from the root of the spanning tree as shown in Figure 4. In block 120 a second set of CC cells are generated and dispatched from the node at

which CC cells are received in block 106 toward the root. The second set of CC cells may be sent at the same or a similar rate to the first set of CC cells. In block 126 the CC cells of the second set of are detected at an end of the connection. The second set of CC cells may be detected, for example, at the same node at which the first set of CC cells originates. As long as the CC cells of the second set of CC cells continue to be detected at the end of the connection it is known that the connection is transmitting cells. The time since receipt of the last CC cell is monitored in block 127. If the time exceeds a threshold as determined in block 128 then an alarm condition is initiated in block 130. Upon the alarm condition being triggered, an action request may optionally be generated (block 132). The action request may comprise, for example, a topology change request or a connection rerouting request. Where a VLAN comprises a plurality of connections extending through one or more cell relay networks, method 100 may be applied to any or all of the connections.

[0039] Method 100 provides a method for detecting a broken leaf-to-root path even in cases where a corresponding root-to-leaf path is not broken. Method 100 further provides a way to avoiding the deleterious effects of broken leaf-to-root paths.

[0040] Detecting the failure of a connection may be performed in various ways. For example, the method described above checks to determine if the time since a last CC cell was received exceeds a threshold time. The threshold time may be chosen such that an alarm condition is triggered only if N sequential CC cells are not received at the location at which block 106 is performed. N is a number which may be, for example, in the range of 1 to 10, and is preferably in the range of 3 to 6.

[0041] Another method for detecting the failure of a connection is to maintain a count of the number of missed CC cells within a window having a given duration. If more than a threshold number of CC cells have been missed within the window then the alarm condition may be triggered. For example, where CC cells are sent at a rate of 1 per second to a cell sink **22** and all of the CC cells are received then cell sink would receive 5 or 6 cells within a 5 ½ second window. If the number of cells received within the window falls below a threshold then the alarm condition may be triggered.

[0042] Those skilled in the art will appreciate that the methods and apparatus of this invention may be applied to internetworking bridged LAN fault tolerance mechanisms with the fault tolerance mechanisms provided by a connection-based network over which data in the bridged LAN is carried.

[0043] Certain implementations of the invention comprise computer processors which execute software instructions which cause the processors to perform a method of the invention. The invention may also be provided in the form of a program product. The program product may comprise any medium which carries a set of computer-readable signals comprising instructions which, when executed by a computer processor, cause the data processor to execute a method of the invention. The program product may be in any of a wide variety of forms. The program product may comprise, for example, physical media such as magnetic data storage media including floppy diskettes, hard disk drives, optical data storage media including CD ROMs, DVDs, electronic data storage media including ROMs, flash RAM, or the like or transmission-type media such as digital or analog communication links.

[0044] Where a component (e.g. a software module, processor, assembly, device, circuit, etc.) is referred to above, unless otherwise indicated, reference

to that component (including a reference to a "means") should be interpreted as including as equivalents of that component any component which performs the function of the described component (i.e., that is functionally equivalent), including components which are not structurally equivalent to the disclosed structure which performs the function in the illustrated exemplary embodiments of the invention.

[0045] As will be apparent to those skilled in the art in the light of the foregoing disclosure, many alterations and modifications are possible in the practice of this invention without departing from the spirit or scope thereof. For example:

- While it is convenient in an ATM network to format the CC packets as OAM cells, other types of cells could be used in place of the OAM cells;
- where the connection-based network comprises another type of network such as an MPLS network, the CC packets may comprise frames sent between nodes along a path along which data for a VLAN is carried in the connection-based network.

Accordingly, the scope of the invention is to be construed in accordance with the substance defined by the following claims.